BOEING SCIENTIFIC RESEARCH LABORATORIES

Further Statistical Problems of Tracking a Target from an Observatory Satellite

Sam C. Saunders

Mathematics Research

April 1965

FURTHER STATISTICAL PROBLEMS OF TRACKING A TARGET

FROM AN OBSERVATORY SATELLITE

by

Sam C. Saunders

SUMMARY

The purpose of this note is to study the statistical problems
of estimation and confidence for both the location and speed and
heading of a target following a linear course when tracked by radar
from an Observatory Satellite (O. S.).

Section I of this report specifies in detail the assumptions
upon which the analysis of the subsequent sections is based. The
main mathematical results are contained in Sections II and III.
The join' distribution of the estimates for target position is derived
in Section II and in Section III the maximum likelihood estimate of
targ t position is obtained and a proof given that it is normally
d' tributed.

Section IV contains a detailed comparison of the mathematical
models of this and the earlier report, and the important theorems and
formulas relating to each. Sections V and VI present special cases
in which the results can be expressed in a form which is simpler than
the general results of Sections II and III.

The last two sections deal respectively with confidence regions
for target speed and heading and the possibilities for tracking a
target in a dense environment given certain information on the
distribution of target speeds and directions.

The notation used in this report is such that much of the detailed
analysis of the previous work Reference [2] applies directly with at

most a reinterpretation of the symbols involved. Those results from the earlier report which do not apply directly to this model are proved to hold for this case.

I. INTRODUCTION

This paper is a continuation of an earlier report, Reference [2], in which the same problem for the location of a target was considered under slightly different assumptions which we call Model I. Both reports make similar assumptions regarding the errors arising from the surveillance of a target by an O. S. The difference lies in the statistical errors made in the determination of the O. S. position from the ground.

In the first model, the position of the O. S. was determined independently for each observation of the target. This would be the case when the observations were made from different O. S.'s or when observations from a single O. S. were so spaced in time that independent assessments of its position were available for each target observation.

In the model presented here, a determination of the orbit of the O. S. is made and its position at a given time estimated. Positions at a later time in the same orbit are then calculated from this estimate. This would apply to situations where several observations are made on a single overflight and no new estimate of the O. S. position is made between the first and last observation. Thus, successive estimates of target position via O. S. observations are independent in the previous Model I while they are dependent in Model II presented here.

As in the first report on this subject, we need to specify precisely the assumptions on which the analysis is based. These are:

1° The O. S. moves at a known constant altitude linearly over the plane with a known constant velocity.

2° The target moves in a straight line at a constant but unknown velocity in the plane.

3° The estimated position coordinates of the O. S. over the plane as determined from the ground at a given time are bivariate normal random variables with known covariance.

4° The estimated position coordinates of the target on the plane as determined relative to the true position of the O. S. are bivariate normal random variables and successive observations of such relative positions are independent.

5° Time between successive observations can be measured with sufficient accuracy so that errors of position due to time inaccuracy are negligible.

6° The parameters of the covariance matrix of the observations of target position relative to the true O. S. position can be determined from bearing angle and range data.

Thus, it is clear that Model II is intended to supplement Model I and not to replace it.

In general, much of the detailed analysis of specific cases
applies directly to both models since we have made a special effort
to utilize the same basic notations in both reports. Also, in
general we try to make the same notational conventions of Reference [2]
apply here. Some of these are:

(i)   Summations will only be over the indices i,j,k,ℓ
      and extend from 1 to n.

(ii) Greek letters, upper and lower case, will in general
      denote respectively random or non-random points in
      the plane as vectors.

(iii) Upper case Roman letters with or without affixes such
      as A,B,... will be matrices.

(iv) Random variables will be indicated by upper case letters
      and by lower case letters with circumflexes.

(v)   Numbered equations of the previous Model I in Reference
      [2] when referred to will be prefixed by a I.

II. OBSERVATIONS OF TARGET POSITION

On a single overflight, an O. S. may make several observations of
the position of a target and our first problem is to determine the
joint distribution of the observations of the target position using the
fact that the target position relative to the O. S. is subject to
observational error as is the estimate of the O. S. positions relative
to the ground. Thus, at a given time t_i we assume the O. S. is at

some position, say ρ_1, and the target at some position μ_1 with respect to a given coordinate system. However, the position of the target as observed by radar from the O. S. is subject to chance error and hence, the radar estimate of the target position from the O. S. position is a random variable, say Ξ_1. Now a radar measurement from the ground at time t_0 of the O. S. position ρ_0 on the given coordinate system is also a random variable, call it Λ.

From assumptions 1° and 2°, we have that the O. S. follows a linear path in the plane, say,

$$\rho_t = \iota + \epsilon t$$

as does the target

$$\mu_t = \alpha + \beta t.$$

Without loss of generality, we can select our coordinates so that the first coordinate of ϵ is in the direction of O. S. travel and hence, the second coordinate of ϵ is zero. Moreover, by assumption 1°, the first coordinate of ϵ is known.

Again, following the general mathematical assumptions 3° and 4° and the notation of our previous model, we have that Ξ_i, for $i = 1, \ldots, n$, and Λ are bivariate normal. More specifically, Ξ_i is a normal variate with mean $\mu_i - \rho_i$ and known covariance matrix C_i^{-1}. That is, Ξ_i is $N(\mu_i - \rho_i, C_i^{-1})$ and thus has density given by

$$\frac{|C_i|^{\frac{1}{2}}}{2\pi} \exp\left\{ -\tfrac{1}{2}(\xi - \mu_i + \rho_i)C_i(\xi - \mu_i + \rho_i)^\dagger \right\} \qquad \xi \in R^2.$$

Also, Λ has mean ρ_0 and covariance matrix D^{-1}, and thus has density

$$\frac{|D|^{\frac{1}{2}}}{2\pi} \exp\left\{-\tfrac{1}{2}(\lambda - \rho_0)D(\lambda - \rho_0)^+\right\} \quad \lambda \in R^2.$$

The data that we obtain from the O. S. yields the observations $\P_i = \Xi_i + \Lambda$ for $i = 1,\ldots,n$ and we seek the joint density of $\P = (\P_1,\ldots,\P_n)$, call it f. Now the joint density of $(\Xi_1,\ldots,\Xi_n,\Lambda)$ is

$$\prod_{i=1}^{n} \frac{|C_i|^{\frac{1}{2}}}{2\pi} \exp\left\{-\tfrac{1}{2}(\xi_i - \mu_i + \rho_i)C_i(\xi_i - \mu_i + \rho_i)^+\right\}\frac{D^{\frac{1}{2}}}{2\pi} \exp\left\{-\tfrac{1}{2}(\lambda - \rho_0)D(\lambda - \rho_0)^+\right\}$$

from which the density f may be found by using

$$\eta_i = \xi_i + \lambda, \quad \eta = (\eta_1,\ldots,\eta_n).$$

Thus,

$$f(\eta) = \int \prod_{1}^{n} \frac{|C_i|^{\frac{1}{2}}}{(2\pi)^n} \exp\left\{-\tfrac{1}{2}\Sigma(\xi_i - \lambda - \mu_i + \rho_i)C_i(\xi_i - \lambda - \mu_i + \rho_i)^+\right\}$$

$$\cdot \frac{|D|^{\frac{1}{2}}}{2\pi} \exp\left\{-\tfrac{1}{2}(\lambda - \rho_0)D(\lambda - \rho_0)^+\right\} d\lambda$$

where the integration extends over the λ-plane. If we denote $\omega = \lambda - \rho_0$, $\zeta_i = \eta_i - \mu_i + \rho_i - \rho_0$ then

$$f(\eta) = \int \prod_{i=1}^{n} \frac{|C_i|^{\frac{1}{2}}|D|^{\frac{1}{2}}}{(2\pi)^{n+1}} \exp\left\{-\tfrac{1}{2}[\Sigma(\zeta_i-\omega)C_i(\zeta_i-\omega)^+ + \omega D\omega^+]\right\} d\omega .$$

Now in the exponent we have

$$[\cdots] = \omega D \omega^{+} + \Sigma[\omega C_i \omega^{+} - \zeta_i C_i \omega^{+} - \omega C_i \zeta_i^{+} + \zeta_i C_i \zeta_i^{+}]$$

$$= \omega(D + \Sigma C_i)\omega^{+} - \Sigma\zeta_i C_i \; \omega^{+} - \Sigma\omega C_i \zeta_i^{+} + \Sigma\zeta_i C_i \zeta_i^{+}.$$

Since for any vector τ and any matrix A we have

$$(\omega-\tau)A(\omega-\tau)^{+} = \omega A \omega^{+} - \tau A \omega^{+} - \omega A \tau^{+} + \tau A \tau^{+},$$

if we let

$$(2.1) \qquad A = D + \Sigma C_i, \qquad \tau = \Sigma\zeta_i C_i A^{-1}$$

then

$$[\cdots] = (\omega - \tau)A(\omega - \tau)^{+} - \tau A \tau^{+} + \Sigma\zeta_i C_i \zeta_i^{+}$$

and

$$f(n) = \frac{|D|^{\frac{1}{2}}\prod_{1}^{n}|C_i|^{\frac{1}{2}}}{(2\pi)^{n}} \exp\left\{-\tfrac{1}{2}[-\tau A \tau^{+} + \Sigma\zeta_i C_i \zeta_i^{+}]\right\}$$

$$\cdot \frac{1}{|A|^{\frac{1}{2}}} \int \frac{|A|^{\frac{1}{2}}}{2\pi} \exp[-\tfrac{1}{2}(\omega - \tau)A(\omega - \tau)^{+}]d\omega.$$

Writing $\zeta = (\zeta_1,\ldots,\zeta_n)$, we now want to find a matrix E so that

$$\zeta E \zeta^{+} = \Sigma\zeta_i C_i \zeta_i^{+} - \tau A \tau^{+}.$$

Substituting (2.1) and expanding both sides, we have

$$\sum_{ij} \zeta_i E_{ij} \zeta_j^{+} = \sum_{i} \zeta_i C_i \zeta_i^{+} - \sum_{ij} \zeta_i C_i A^{-1} C_j \zeta_j^{+},$$

using the fact that both D and C, hence A, are symmetric. Thus, we have

$$
(2.2) \qquad E_{ij} = \begin{cases} -C_i A^{-1} C_j & \text{if} \quad i \neq j \\ \\ C_i - C_i A^{-1} C_i & \text{if} \quad i = j, \end{cases}
$$

from which one checks that $E_{ij} = E_{ji}^{\dagger}$ and hence that E is symmetric, itself. Denoting $\upsilon_i = \mu_i - \rho_i + \rho_0$ then $\zeta_i = \eta_i - \upsilon_i$ and $\zeta = \eta - \upsilon$, we have the fundamental

THEOREM 1: The distribution of the observations of target position $\eta = (\eta_1, \ldots, \eta_n)$, obtained from the O. S. is

$$
(2.3) \qquad f(\eta) = \frac{|E|^{\frac{1}{2}}}{(2\pi)^n} \exp\left[-\tfrac{1}{2}(\eta - \upsilon)E(\eta - \upsilon)^{\dagger}\right] \qquad \eta \in R^{2n}
$$

where

$$
(2.4) \qquad |E| = \frac{|D|\prod_{1}^{n} |C_i|}{|A|}.
$$

Thus, we conclude that η is distributed as $N(\upsilon, E^{-1})$. That is, η is normally distributed with mean υ and covariance matrix E^{-1}.

To complete the proof of Theorem 1, we should verify Relation (2.4). From (2.2) we have

$$|E| = \begin{vmatrix} C_1 - C_1 A^{-1} C_1 & - C_2 A^{-1} C_1 & - C_3 A^{-1} C_1 & \cdots & - C_n A^{-1} C_1 \\ - C_1 A^{-1} C_2 & C_2 - C_2 A^{-1} C_2 & - C_3 A^{-1} C_2 & \cdots & - C_n A^{-1} C_2 \\ - C_1 A^{-1} C_3 & - C_2 A^{-1} C_3 & C_3 - C_3 A^{-1} C_3 & \cdots & - C_n A^{-1} C_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ - C_1 A^{-1} C_n & - C_2 A^{-1} C_n & - C_3 A^{-1} C_n & \cdots & C_n - C_n A^{-1} C_n \end{vmatrix}.$$

Now we may factor C_i from the ith column for each $i = 1, \ldots, n$ and subtract the first column from the 2nd, $3^{rd}, \ldots$, last column to obtain

$$|E| = \prod_{i=1}^{n} |C_i| \begin{vmatrix} I - A^{-1} C_1 & -I & -I \cdots -I \\ - A^{-1} C_2 & I & 0 \cdots 0 \\ - A^{-1} C_3 & 0 & I \cdots 0 \\ \vdots & \vdots & \vdots \vdots \\ - A^{-1} C_n & 0 & 0 \cdots I \end{vmatrix}.$$

Multiplying the ith column by $A^{-1} C_i$ and adding to the first column, we get

$$\frac{|E|}{\prod |C_i|} = \begin{vmatrix} I - \iota A^{-1} C_i & -I & -I \cdots -I \\ 0 & I & 0 \cdots 0 \\ 0 & 0 & I \cdots 0 \\ \vdots & \vdots & \vdots \cdots \vdots \\ 0 & 0 & 0 \cdots I \end{vmatrix} = |I - \Sigma A^{-1} C_i|.$$

To complete our check, we show that

$$|I - \Sigma A^{-1}C_i| = |A^{-1}A - \Sigma A^{-1}C_i| = |A^{-1}||A - \Sigma C_i| = \frac{|D|}{|A|} .$$

III. THE MAXIMUM LIKELIHOOD ESTIMATES AND THEIR DISTRIBUTION

Theorem 1 tells us that a single observation of O. S. position from a bivariate normal distribution relative to the ground combined with n observations of target position which are normally distributed relative to the O. S. position yields a joint normal distribution of observed target positions relative to the ground. Our next problem is to find a best estimate of target course using this normal distribution of error and the assumption that the target is moving linearly at constant velocity.

Thus, we want to find the maximum likelihood estimate (m.l.e.) of future target positions μ_t, call the estimate $\hat{\mu}_t$. Now by the known properties of the maximum likelihood estimates we know that

$$\hat{\mu}_t = \hat{\alpha} + \hat{\beta}t,$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the m.l.e. of α and β and to complete the picture we need to know the distribution of $(\hat{\alpha}, \hat{\beta})$.

Now we see that

$$\upsilon_i = \mu_i - \rho_i + \rho_0 = (\alpha + \varepsilon t_0) + (\beta - \varepsilon)t_i.$$

If we define

$$\gamma = \alpha + \epsilon t_0, \qquad \kappa = \beta - \epsilon$$

we have a model similar to that analysed previously and if we obtain the maximum likelihood estimates of γ and κ, then we can easily obtain the m.l.e. of α, β.

The likelihood function is

$$L = \text{Ln}|E| - n\text{Ln}(2\pi) - \sum_{ij} (\eta_i - \upsilon_i) E_{ij} (\eta_j - \upsilon_j)^\dagger.$$

Let $F = \zeta_i E_{ij} \zeta_j^\dagger$ with $\zeta_i = \eta_i - \upsilon_i$ and let us for this argument introduce the notation for the coordinates of

$$\gamma = (\gamma_1, \gamma_2), \qquad \kappa = (\kappa_1, \kappa_2)$$

then

$$\left.\begin{aligned}
\frac{\partial F}{\partial \gamma_k} &= \zeta_i E_{ij} \left(\frac{\partial \zeta_j}{\partial \gamma_k}\right)^\dagger + \left(\frac{\partial \zeta_i}{\partial \gamma_k}\right) E_{ij} \zeta_j^\dagger \\
\frac{\partial F}{\partial \kappa_k} &= \zeta_j E_{ij} \left(\frac{\partial \zeta_j}{\partial \kappa_k}\right)^\dagger + \left(\frac{\partial \zeta_i}{\partial \kappa_k}\right) E_{ij} \zeta_j^\dagger
\end{aligned}\right\} \qquad k = 1,2$$

Since

$$\frac{\partial \zeta_j}{\partial \gamma_k} = \frac{\partial (\eta_j - \gamma - \kappa t_j)}{\partial \gamma_k} = -(\delta_{1k}, \delta_{2k}) = -\delta_k$$

$$\frac{\partial \zeta_j}{\partial \kappa_k} = \frac{\partial (\eta_j - \gamma - \kappa t_j)}{\partial \kappa_k} = -t_j (\delta_{1k}, \delta_{2k}) = -t_j \delta_k$$

holds for all $j = 1,\ldots,n$ with δ_{ij} the Kronecker delta, we obtain

$$(3.1) \qquad \frac{\partial L}{\partial \gamma_k} = \frac{1}{2} \sum_{ij} (\zeta_i E_{ij} \delta_k^\dagger + \delta_k E_{ij} \zeta_j^\dagger)$$

$$k = 1,2$$

$$(3.2) \qquad \frac{\partial L}{\partial \kappa_k} = \frac{1}{2} \sum_{ij} (t_j \zeta_i E_{ij} \delta_k^\dagger + t_i \delta_k E_{ij} \zeta_j^\dagger).$$

Thus, we have from (3.1), by realizing that the second term is a scalar and thus equal to its own transpose,

$$\sum_{ij} (\zeta_i + \zeta_j) E_{ij} \delta_k^\dagger = 0 \qquad\qquad k = 1,2,$$

or writing these two equations in matrix form

$$\sum_{ij} (\zeta_i + \zeta_j) E_{ij} (\delta_1^\dagger, \delta_2^\dagger) = (0,0).$$

But since $(\delta_1^\dagger, \delta_2^\dagger) = I$, (3.1) reduces to

$$(3.3) \qquad \sum_{ij} (\zeta_i + \zeta_j) E_{ij} = (0,0).$$

Now from (3.2), following similar steps, we have

$$\sum_{ij} (t_j \zeta_i + t_i \zeta_j) E_{ij} \delta_k^\dagger = 0 \qquad\qquad k = 1,2$$

from which we obtain

$$(3.4) \qquad \sum_{ij} (t_j \zeta_i + t_i \zeta_j) E_{ij} = (0,0).$$

But we note

$$\zeta_i + \zeta_j = \eta_i - \gamma - \kappa t_i + \eta_j - \gamma - \kappa t_j = \eta_i + \eta_j - 2\gamma - \kappa(t_i + t_j)$$

$$t_j\zeta_i + t_i\zeta_j = t_j\eta_i - \gamma t_j - \kappa t_i t_j + t_i\eta_j - \gamma t_i - \kappa t_i t_j$$

$$= t_j\eta_i + t_i\eta_j - \gamma(t_i + t_j) - 2\kappa t_i t_j.$$

Hence by substitution, we finally obtain from (3.1) the equation

$$(3.5) \qquad \sum_{ij}(\eta_i + \eta_j)E_{ij} = 2\gamma\sum_{ij}E_{ij} + \kappa\sum_{ij}(t_i + t_j)E_{ij}$$

and from (3.2)

$$(3.6) \qquad \sum_{ij}(t_j\zeta_i + t_i\eta_j)E_{ij} = \gamma\sum_{ij}(t_i + t_j)E_{ij} + 2\kappa\sum_{ij}t_i t_j E_{ij}.$$

Writing these in matrix notation, we have

$$(3.7) \qquad (\gamma,\kappa)\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21}, & \Sigma_{22} \end{pmatrix} = (\psi_1,\psi_2),$$

where

$$\Sigma_{11} = \sum_{ij}E_{ij}$$

$$\Sigma_{12} = \Sigma_{21} = \sum_{ij}\frac{t_i + t_j}{2}E_{ij}$$

$$\Sigma_{22} = \sum_{ij}t_i t_j E_{ij}.$$

Now if we define the matrix $F = (F_{ij})$ by

$$F_{ij} = \frac{E_{ij} + E_{ji}}{2}$$

which we write in matrix form $F = \dfrac{E + E^\circ}{2}$ with the obvious definition

$$E^\circ = (E^\circ_{ij}) \quad \text{where} \quad E^\circ_{ij} = E_{ji},$$

then we can see, by rearranging some summations,

$$(3.8.1) \qquad \Sigma_{11} = \sum_{ij} F_{ij}$$

$$(3.8.2) \qquad \Sigma_{12} = \Sigma_{21} = \sum_{ij} t_i F_{ij}$$

$$(3.8.3) \qquad \Sigma_{22} = \sum_{ij} t_i t_j F_{ij}$$

and in the same manner

$$(3.8.4) \qquad \cdot_1 = \sum_{ij} \frac{r_i + r_j}{2} E_{ij} = \sum_{ij} r_i F_{ij}$$

$$(3.8.5) \qquad \cdot_2 = \sum_{ij} \frac{t_j r_i + t_i r_j}{2} E_{ij} = \sum_{ij} r_i t_j F_{ij}.$$

Now if we denote

$$(3.8.6) \qquad S^{-1} = (\Sigma_{ij}) \qquad S = (S_{ij}) \qquad i,j = 1,2$$

then (3.7) may be written $(\hat{\gamma}, \cdot) = (\psi_1, \psi_2)S$, or

(3.9.1) $\hat{\gamma} = \psi_1 S_{11} + \psi_2 S_{21}$

(3.9.2) $\cdot = \psi_1 S_{12} + \psi_2 S_{22}$

where, of course, $S_{21} = S_{12}^{\dagger}$. We now write from (3.8.4), (3.8.5)

$$(\psi_1, \psi_2) = (\eta_1, \ldots, \eta_n) \begin{pmatrix} \sum\limits_j F_{1j}, & \sum\limits_j t_j F_{1j} \\[2em] \sum\limits_j F_{nj}, & \sum\limits_j t_j F_{nj} \end{pmatrix} = \eta U$$

with η, U defined in the obvious manner. Since by Theorem 1,
η is $N(\nu, E^{-1})$ and linear transformations of normal variables
are again normal, it is clear that $\cdot = (\psi_1, \psi_2)$ is bivariate
normal. To be precise, by P. 19 of Reference [1], we have that ψ
is $N(\nu U, U^{\dagger}E^{-1}U)$. Since $(\hat{\gamma}, \cdot) = \cdot S$, we know by a similar
argument that $(\hat{\nu}, \cdot)$ is $N(\nu US, Q)$ where

(3.9.3) $Q = S^{\dagger}U^{\dagger}E^{-1}US.$

We will now prove that

(3.10) $\nu US = (\nu, \cdot).$

To do this, we make use of the identities which follow from
$S^{-1}S = I$

$$\Sigma_{11}S_{11} + \Sigma_{12}S_{21} = I, \qquad \Sigma_{11}S_{12} + \Sigma_{12}S_{22} = 0$$

(3.10.01)

$$\Sigma_{21}S_{11} + \Sigma_{22}S_{21} = 0, \qquad \Sigma_{21}S_{12} + \Sigma_{22}S_{22} = I.$$

The first component of the vector νUS is

$$\sum_{ij}(\gamma + \kappa t_i)F_{ij}S_{11} + \sum_{ij}(\gamma + \kappa t_i)t_jF_{ij}S_{21} =$$

$$= \gamma(\Sigma_{11}S_{11} + \Sigma_{12}S_{2.}) + \kappa(\Sigma_{21}S_{11} + \Sigma_{22}S_{21}) = \gamma.$$

The second component of the vector νUS is

$$\sum_{ij}(\gamma + \kappa t_i)F_{ij}S_{12} + \sum_{ij}(\nu + \kappa t_i)t_jF_{ij}S_{22} =$$

$$= \gamma(\Sigma_{11}S_{12} + \Sigma_{12}S_{22}) + \kappa(\Sigma_{21}S_{12} + \Sigma_{22}S_{22}) = \kappa.$$

This proves (3.10). Now under the additional assumption that

(3.10.1) $$E^\circ E^{-1} E^\circ = E$$

we can show that

(3.10.2) $$Q = S.$$

If we write

$$U = FT^\circ = \begin{pmatrix} F_{11} & \cdots & F_{1n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ F_{n1} & \cdots & F_{nn} \end{pmatrix} \begin{pmatrix} I, & t_1 I \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ I, & t_n I \end{pmatrix}$$

where

$$(3.10.3) \qquad T_j = (I, t_j I) \qquad T = (T_1, \ldots, T_n)$$

then

$$U^\dagger E^{-1} U = TFE^{-1} FT^\circ$$

since $F^\dagger = F$ we have

$$4FE^{-1}F = (E + E^\circ)E^{-1}(E + E^\circ)$$

$$= E + E^\circ + E^\circ + E^\circ E^{-1} E^\circ = 2F.$$

Now

$$TFT^\circ = \sum_{ij} \begin{pmatrix} F_{ij} & t_j F_{ij} \\ t_i F_{ij} & t_i t_j F_{ij} \end{pmatrix} = S^{-1}$$

and the assertion follows since $S^\dagger = S$.

We have just obtained

THEOREM 2: The maximum likelihood estimates $(\hat{\alpha}, \hat{\beta})$ defined

by $\qquad \hat{\alpha} = \hat{\gamma} - t_0, \qquad \hat{\beta} = \hat{\kappa} + t$

where $(\hat{\gamma}, \hat{\kappa})$ are defined by (3.9.1), (3.9.2) and are $N((\alpha, \beta), Q)$

with Q given in (3.9.3).

From this we have the immediate

COROLLARY 3: The maximum likelihood estimate of the true target

position u_t at time t is

$$\hat{\mu}_t = \hat{a} + \hat{\beta}t$$

which is distributed as $N(\mu_t, B_t)$

where

(3.11) $\qquad B_t = Q_{11} + t(Q_{12} + Q_{21}) + t^2 Q_{22}.$

We will now prove

THEOREM 3: The estimate $\hat{\mu}_t$ and its covariance matrix B_t are invariant under location and scale change in time.

PROOF: Let $t' = at + b$ be a linear transformation of time and use primes affixed to matrices to denote that matrix computed at the new times

(3.12) $\quad \Sigma'_{11} = \Sigma_{11}, \qquad \Sigma'_{21} = a\Sigma_{21} + b\Sigma_{11}, \qquad \Sigma'_{22} = a^2\Sigma_{22} + 2ab\Sigma_{21} + b^2\Sigma_{11}.$

Now

$$\hat{\mu}_t = \hat{a} + \hat{\beta}t = \hat{\gamma} + \hat{\kappa}t = \epsilon(t - t_0)$$

where ϵ represents the velocity of the O. S. per unit of time. It is thus sufficient to show that $\hat{\gamma} + \hat{\kappa}t$ is invariant. In our previous notation

$$(\hat{\gamma}, \hat{\kappa}) = nFT^\circ S$$

where nF is independent of time. Hence

$$\hat{\gamma} + \hat{\kappa}t = F \begin{pmatrix} S_{11} + t_1 S_{21} + tS_{12} + tt_1 S_{22} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ S_{11} + t_n S_{21} + tS_{12} + tt_n S_{22} \end{pmatrix} \quad .$$

It is sufficient to show that for each $j = 1,\ldots,n$ that

$$(3.13) \qquad T_j S T_i^\dagger = S_{11} + t_j S_{21} + t_i S_{12} + t_i t_j S_{22}$$

is invariant.

From the equations (3.10.01), (3.10.02), we have

$$S_{22}^{-1} = \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$$

$$S_{12} = -\Sigma_{11}^{-1} \Sigma_{12} S_{22} \qquad\qquad S_{21} = S_{12}^\dagger$$

$$S_{11} = \Sigma_{11}^{-1}(I - \Sigma_{12} S_{21}).$$

There follows by substitution from (3.1.2)

$$S_{22}'^{-1} = a^2(\Sigma_{22} - \Sigma_{21}\,{}_{11}^{-1}\Sigma_{12}) = a^2 S_{22}^{-1}$$

hence

$$S_{22}' = \frac{1}{a^2} S_{22}. \quad \text{Likewise, we find}$$

$$S_{12}' = \frac{S_{12}}{a} - \frac{b}{a} S_{22}, \qquad S_{21}' = \frac{1}{a} S_{21} - \frac{b}{a^2} S_{22}$$

$$S_{11}' = S_{11} - \frac{b}{a}(S_{21} + S_{12}) + \frac{b^2}{a^2} S_{22}.$$

Now we calculate

$$T_j' S' T_i'^\dagger = S_{11} - \frac{b}{a}(S_{21} + S_{12}) + \frac{b^2}{a^2} S_{22} + (at_j + b)(\frac{1}{a} S_{21} - \frac{b}{a^2} S_{22})$$

$$+ (at_i + b)(\frac{1}{a} S_{12} - \frac{b}{a^2} S_{22}) + (at_i + b)(at_j + b) \frac{1}{a^2} S_{22}$$

expanding and simplifying shows

$$= S_{11} + t_j S_{21} + t_i S_{12} + t_i t_j S_{22} = T_j ST_i .$$

Thus, we have shown that $\hat{a} + \hat{s}t$ is invariant under scale and location change in time.

To see that B_t is also invariant under the same transformation of time we merely note that if we write $T_0 = (1, t_0 1)$ then by making use of the notation of (3.10.3) and the fact that $U = FT^\circ$ we have

$$B_{t_0} = T_0 Q T_0 = (T_0 ST) F E^{-1} F (TST_0)$$

but $T_0 ST$ is itself invariant by our previous argument.

IV. A COMPARISON OF MODELS I AND II

We have proceeded thus far without a direct comparison with Model I. As in Section II of Reference [2], we assume that at a given time t_i we have the O. S. at point c_i
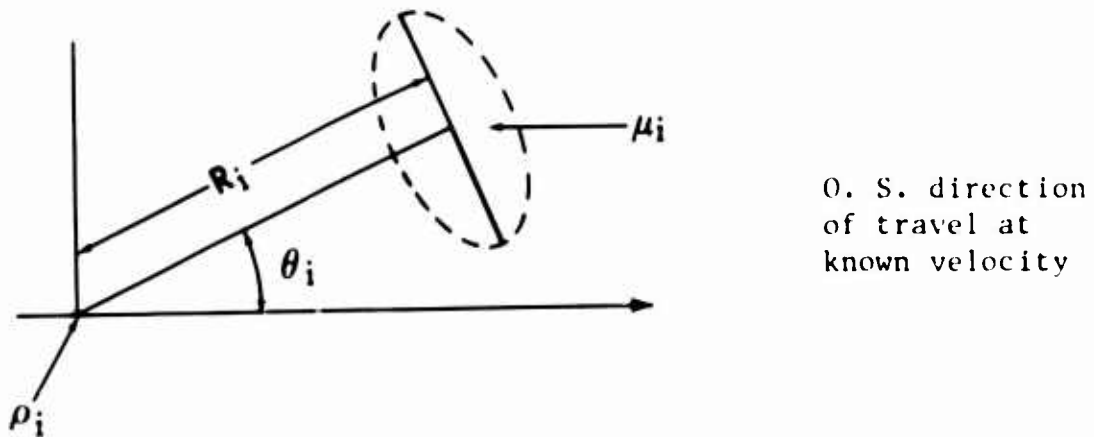

O. S. direction
of travel at
known velocity

FIGURE 1

observing, at bearing angle θ_i and range R_i, the target position μ_i. The observed position of the target is the random variable z_i which by assumption is $N(\mu_i - \mu_i, C_i^{-1})$. By following the same argument used in Section II of Reference [2], we have

$$(4.1) \qquad C_i^{-1} = \begin{pmatrix} \sigma^2_{x_i} & \sigma_{x_i y_i} \\ \\ \sigma_{x_i y_i} & \sigma^2_{y_i} \end{pmatrix}$$

where exactly as (1.2.3) we have

$$(4.2) \qquad \sigma_{x_i y_i} = \text{cov}(x_i y_i) = \frac{\sigma^2_{u_i} - \sigma^2_{v_i}}{2} \sin 2\theta_i.$$

Following (1.7.1.2) and (1.7.1.3) we have

$$(4.3) \qquad \sigma^2_{x_i} = \sigma^2_{u_i} \cos^2 \theta_i + \sigma^2_{v_i} \sin^2 \theta_i$$

$$(4.4) \qquad \sigma^2_{y_i} = \sigma^2_{u_i} \sin^2 \theta_i + \sigma^2_{v_i} \cos^2 \theta_i.$$

Now at time t_i, σ_{u_i} is the standard deviation of the range error and σ_{v_i} is the standard deviation of the azimuth error and both are known functions of the given operational range R_i. It is clear that we also need the assumption $6°$ that the parameters of C_i can be determined by the bearing angle θ_i and the range R_i.

Now we also assume as in Model I that the matrix D is diagonal and known, namely

$$D^{-1} = \begin{pmatrix} \sigma_z^2 & & 0 \\ & & \\ 0 & & \sigma_w^2 \end{pmatrix} .$$

The difference in the models is essentially in the different densities of the set of observations $\eta = (\eta_1, \ldots, \eta_n)$.

In Model I we have

$$f(\eta) = \frac{\prod_{i=1}^{n} |A_i|^{\frac{1}{2}}}{(2\pi)^n} \exp\left\{ -\tfrac{1}{2} \sum_{1}^{k} (\eta_i - \mu_i) A_i (\eta_i - \mu_i)^+ \right\}$$

where $A_i^{-1} = C_i^{-1} + D^{-1}$ for $i = 1, \ldots, n.$

On the other hand, the density of η for Model II is given in Theorem i, Equation (1.10), which may be compared with that above.

V. SPECIAL CASE OF THE DISTRIBUTION OF RADAR ERRORS CONSTANT IN TIME

For this case, we assume that neither the orientation of the O. S.'s direction of travel nor its position relative to the target plays a role in determining the covariance matrix of the observation; thus, we have

$$C_i = C \qquad \text{for} \qquad i = 1, \ldots, n.$$

Now from (2.1) we have

$$A = D + nC$$

and from (2.2)

$$E_{ij} = \delta_{ij}C - CA^{-1}C = E_{ji}.$$

Then $Q = S$ and $F_{ij} = E_{ij}$

$$\Sigma_{11} = \sum_{ij} E_{ij} = \sum_i (C - nCA^{-1}C) = nC - n^2CA^{-1}C,$$

$$\Sigma_{12} = \Sigma_{21} = \sum_{ij} t_i E_{ij} = \sum_i (t_iC - nt_iCA^{-1}C) = n\bar{t}C - n^2\bar{t}CA^{-1}C,$$

$$\Sigma_{22} = \sum_{ij} t_i t_j E_{ij} = \sum_i (t_i\delta_{ij}C - t_jCA^{-1}C),$$

$$= \sum_i t_i(t_jC - n\bar{t}CA^{-1}C) = n\overline{t^2}C - (n\bar{t})^2CA^{-1}C,$$

$$\psi_1 = \sum_{ij} \eta_i E_{ij} = \sum_i (\eta_iC - n\bar{\eta}CA^{-1}C) = n\bar{\eta}(C - nCA^{-1}C),$$

$$\psi_2 = \sum_{ij} \eta_i t_j E_{ij} = \sum_i \eta_i(t_iC - n\bar{t}CA^{-1}C) = n\overline{\eta t}C - n^2\bar{t}CA^{-1}C.$$

By the invariance theorem, we can choose the time origin
so that $\bar{t} = 0$. Hence, we have

$$\Sigma_{11} = nC - n^2 CA^{-1}C,$$

(5.0.8)
$$\Sigma_{21} = C = \Sigma_{12}$$

$$\Sigma_{22} = n\overline{t^2}C,$$

$$\psi_1 = \overline{n}\Sigma_{11}, \qquad \psi_2 = n\overline{t\eta}C.$$

Now

$$S_{11} = \Sigma_{11}^{-1}, \qquad S_{12} = 0 = S_{21}, \qquad S_{22} = \Sigma_{22}^{-1},$$

(5.0.9)
$$B_t = S_{11} + t^2 S_{22} = \Sigma_{11}^{-1} + t^2 \Sigma_{22}^{-1}.$$

and we see from the above

$$\hat{\gamma} = \psi_1 S_{11} = \overline{\eta},$$

$$\hat{\rho} = \psi_2 S_{12} = \frac{\overline{t\eta}}{\overline{t^2}},$$

(5.1) $\quad \hat{\alpha} = \overline{\eta} - \hat{\rho}t_0, \qquad \hat{\beta} = \frac{\overline{t}}{\overline{t^2}} + \cdots$

These two equations may be compared with equations (6.4.1) and
(6.4.2) of Reference [2] with $\overline{t} = 0$.

From (5.0.8) we have

$$\lambda_{11} = nCA^{-1}(A - nC) = nCA^{-1}D,$$

$$\Sigma_{11}^{-1} = \frac{1}{n} D^{-1}AC^{-1} = \frac{1}{n} C^{-1} + D^{-1},$$

$$\Sigma_{22}^{-1} = \frac{1}{nt^2} C^{-1}.$$

Now substituting into (5.0.9), we have

$$(5.2) \qquad B_t = \frac{1}{n} (1 + \frac{t^2}{t^2})C^{-1} + D^{-1}.$$

Thus, for this case we have easy formulas for the estimate $\hat{\mu}_t$ and the covariance matrix B_t.

Of course, if we assume further that C is diagonal, since D is, the calculation of B_t is immediate.

We remark that a comparison may also be made between the two models by examining (5.2) above and (6.3) of Reference [2], remembering that in Reference [2] $A^{-1} = C^{-1} + D^{-1}$.

VI. THE SPECIAL CASE OF SYMMETRIC OBSERVATIONS

In this case we have the n observations symmetrically spaced in time. Using the invariance theorem, we can choose the origin so that we have the relations

$$t_i + t_{n+1-i} = 0$$
$$\qquad\qquad\qquad i = 1,\ldots,n$$
$$\theta_i + \theta_{n+1-i} = \pi$$

for the observation times and the bearing angles. Now if we set

$$a_i = \sigma^2_{x_i}, \qquad b_i = \sigma_{x_i y_i}, \qquad c_i = \sigma^2_{y_i},$$

then

$$C_i^{-1} = \begin{pmatrix} a_i & b_i \\ b_i & c_i \end{pmatrix} \quad \text{and} \quad C_i = \begin{pmatrix} d_i & e_i \\ e_i & f_i \end{pmatrix}.$$

Letting $\Delta_i = |C_i^{-1}| = a_i c_i - b_i^2$ we can write

$$e_i = \frac{-b_i}{\Delta_i}, \qquad f_i = \frac{a_i}{\Delta_i}, \qquad d_i = \frac{c_i}{\Delta_i}$$

and in particular we have

$$e_i = \frac{\sigma^2_{v_i} - \sigma^2_{u_i}}{2\Delta_i} \sin 2\theta_i.$$

From the above relations we see that

$$d_i = d_{n-i+1},$$

$$f_i = f_{n-i+1}, \qquad i = 1,\ldots,n,$$

$$e_i = -e_{n-i+1},$$

so that

$$C_i + C_{n-i+1} = \begin{pmatrix} 2d_i & 0 \\ 0 & 2f_i \end{pmatrix}.$$

Thus we see that the matrix

$$D_1 = \Sigma C_i$$

is diagonal as in $A = D + D_1$. A matrix of the form

$$P = \begin{pmatrix} 0 & P_1 \\ P_2 & 0 \end{pmatrix}$$

we shall call contra-diagonal. Hence, the matrix

$$D_2 = \Sigma t_i C_i$$

is contra-diagonal. Since

$$E_{ij} = \delta_{ij} C_i - C_i A^{-1} C_j,$$

we have

$$\Sigma_{11} = \sum_{ij} E_{ij} = \sum_i (\sum_j \delta_{ij} C_i - \sum_j C_i A^{-1} C_j)$$

$$= \sum_i (C_i - C_i A^{-1} D_1)$$

$$= D_1 - D_1 A^{-1} D_1.$$

Since the inverse product and difference of diagonal matrices are diagonal, we have that Σ_{11} is diagonal. Now

$$\Sigma_{21} = \sum_{ij} t_i E_{ij} = \sum_{ij} t_i (\delta_{ij} C_i - C_i A^{-1} C_j)$$

$$= \sum_i t_i (C_i - t_i C_i A^{-1} D_1) = D_2 - D_2 A^{-1} D_1.$$

And Σ_{22} is diagonal since the product of contra-diagonal matrices is diagonal. Calculation from this point proceeds in a straight-forward manner. Since we have

$$S_{21} = (\Sigma_{21} - \Sigma_{11} \,_{21}^{-1}\Sigma_{22})^{-1},$$

by equations (3.10.01)

(6.4) $\qquad S_1 = -\Sigma_{21}^{-1}\Sigma_{22}S_{21}, \qquad S_{22} = -\Sigma_{21}^{-1}\Sigma_{11}S_{12},$

we can obtain the structure of S. Since Σ_{21} is contra-diagonal so is its inverse, and in particular if

$$\Sigma_{21} = \begin{pmatrix} 0 & d_{12} \\ d_{21} & 0 \end{pmatrix}, \quad \text{then} \quad \Sigma_{21}^{-1} = \begin{pmatrix} 0 & d_{21}^{-1} \\ d_{12}^{-1} & 0 \end{pmatrix}.$$

Since Σ_{21} and Σ_{22} are diagonal, $(S_{21})^{-1}$ is the difference of two contra-diagonal matrices and hence is contra-diagonal. There-fore, we conclude that S_{12} is contra-diagonal. Thus, we see by (6.4) that S_{11} and S_{22} are both diagonal. This, of course, is essentially what we have computed in Section 7 of Reference [1].

Note that in this case in general we do not have Σ_{21} symmetric, as we had for Model I.

Thus, we see that all elements of S are zero except those on the diagonal or contra-diagonal. This knowledge may be of help in the computation of Q for in general for this case we do not have that Q = S.

VII. ESTIMATION AND CONFIDENCE FOR SPEED AND HEADING

The parameters with which one is usually concerned when
tracking a moving ship target are the speed and heading, and in
this section we shall consider the accuracy with which they can
be estimated. In both Models I and II we have obtained the
estimate, $\hat{m}(t) = \hat{a} + \hat{\beta}t$, which we now write in coordinate form

$$\hat{m}(t) = (\hat{m}_1(t), \hat{m}_2(t)).$$

This is an estimate of the target position μ_t at any time t
when it follows a linear path in the plane. Now the true velocity
s of the ship is

$$s = \{[m_1'(t)]^2 + [m_2'(t)]^2\}^{\frac{1}{2}}$$

and the true heading h is the angle

$$\arctan \frac{m_2(1) - m_2(0)}{m_1(1) - m_1(0)}$$

plus a constant depending upon the signs of the numerator and
denominator of the argument of the arctan function. A little more
detailed analysis shows that, since $\beta = (b_1, b_2)$,

$$s = \sqrt{b_1^2 + b_2^2},$$

and

$$h - \arctan \frac{b_2}{b_1} = \begin{cases} 0 & \text{if} \quad b_1 > 0, \ b_2 > 0 \\ \pi & \text{if} \quad b_1 < 0 \\ 2\pi & \text{if} \quad b_1 > 0, \ b_2 < 0 \end{cases}$$

where arctan is the principal value of the inverse tangent and takes values on $(-\pi/2, \pi/2)$.

These relationships are seen to be simply

(7.1) $\qquad b_1 = s \cos h, \qquad b_2 = s \sin h.$

Thus, by analogy we have the equations

(7.2) $\qquad \hat{b}_1 = V \cos \phi, \qquad \hat{b}_2 = V \sin \phi$

defining the random variables V (for velocity) and ϕ for the heading angle, which are estimates of the true speed s and the true heading h. (Here we violate our convention to use Greek letters for vectors in order to retain the more universal one that Greek letters are angles.)

We were able to show that the estimators $(\hat{\alpha}, \hat{\beta})$ are $N(\alpha, \beta), Q)$ (Theorem 1 here and Theorem 2 of Ref. [2]), where

$$Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}$$

is known. Now from known results for normal variates (p. 24 of Reference [1]) it follows that $\hat{\beta}$ is $N(\beta, Q_{22})$.

The joint density of ϕ, V is found by making a transformation to polar coordinates in the probability density of $N(\beta, Q_{22})$. This is

$$(7.3) \quad g(\phi, v | h, s) = \frac{v}{2\pi |Q_{22}|^{\frac{1}{2}}} \exp \{-\tfrac{1}{2} \zeta Q_{22}^{-1} \zeta^{+}\} \qquad \begin{array}{c} 0 < v < \infty \\ \\ 0 < \phi < 2\pi \end{array}$$

where were we have

$$\zeta = (v \cos \phi - s \cos h, \; v \sin \phi - s \sin h).$$

This density can be used to study the distribution of velocity and heading estimates that could arise under special circumstances such as infrequent headings and high velocity.

One might desire separate confidence intervals on the heading and on the velocity. However, if we proceed to find the marginal densities of v and ϕ from (7.3), we see that each density has both parameters h and s.

Thus a confidence interval for the velocity could be constructed only if we knew the true heading h. Likewise, a confidence interval could be found for the true heading if we knew the true velocity. The presence of the nuisance parameters prevents us from obtaining confidence intervals separately when both parameters are unknown.

However, we can obtain a joint confidence region for (h, s), which is somewhat inefficient, as follows. From well-known results on the Chi-square distribution of the quadratic form of normal variates,

we have

$$(7.4) \qquad P[(\hat{\beta} - \beta)Q_{22}^{-1}(\hat{\beta} - \beta)^{+} \leq \chi_2^2(p)] = 1 - p$$

where $\chi_2^2(p)$ was given in Equation (5.2) of Reference [2].
Now if one draws the ellipse centered at $\hat{\beta}$, letting $\zeta = (x,y)$
here be an arbitrary point in the plane, we can specify the ellipse
as

$$(7.5) \qquad W = \{\zeta: (\zeta - \hat{\beta})Q_{22}^{-1}(\zeta - \hat{\beta})^{+} \leq \chi_2^2(p)\}.$$

This defines a $100(1-p)$ per cent confidence region for β, that is,

$$P[\beta \ \epsilon \ W] = 1 - p.$$

Now let us construct the smallest area in polar coordinates which
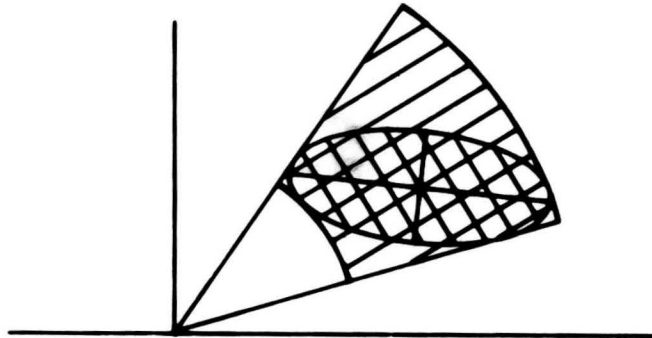is the Cartesian product of intervals and contains W. Let us call
it W^*.

FIGURE 2

If $\zeta = (x,y) \in R^2$, we define $\text{mod}(\zeta) = \sqrt{x^2 + y^2}$, and

$\arg(\zeta) = \arctan(y/x)$. Then we set

(7.5.1) $\qquad \phi_2 = \sup_{\zeta \in W} \arg(\zeta), \quad \phi_1 = \inf_{\zeta \in W} \arg(\zeta)$

(7.5.2) $\qquad V_2 = \sup_{\zeta \in W} \text{mod}(\zeta), \quad V_1 = \inf_{\zeta \in W} \text{mod}(\zeta)$

and define, in turn

(7.5.3) $\qquad W^* = \{(\theta, r) : \phi_1 < \theta < \phi_2, V_1 < r < V_2\}.$

Now $(r \cos \theta, r \sin \theta) \in W$ implies $(\theta, r) \in W^*$. Thus

$P[(h,s) \in W^*] \geq P[(s \cos h, s \sin h) \in W] = P[\beta \in W] = 1 - p.$

The task to which we now address ourselves is the determination

of random variables (ϕ_i, V_i), $i = 1,2$.

The boundary of the elliptical region W given in (7.5) is

(7.6) $\qquad (x - \hat{b}_1, y - \hat{b}_2)Q_{22}^{-1}(x - \hat{b}_1, y - \hat{b}_2)^+ = \chi_2^2(p).$

If we write

$$Q_{22}^{-1} = \begin{pmatrix} s_{11}, & s_{12} \\ s_{21}, & s_{22} \end{pmatrix},$$

by symmetry $s_{12} = s_{21}$, and if we use German letters to denote

the constants defined by

$$a = \frac{s_{11}}{x_2^2(p)} , \qquad b = \frac{s_{12}}{x_2^2(p)} , \qquad c = \frac{s_{22}}{x_2^2(p)}$$

then we may write (7.6) as

$$(7.7) \qquad a(x - \hat{b}_1)^2 + 2b(x - \hat{b}_1)(y - \hat{b}_2) + c(y - \hat{b}_2)^2 = 1.$$

We want to find the maximum and minimum of the functions

$$(7.7.1) \qquad f(x,y) = y/x, \qquad g(x,y) \ x^2 + y^2$$

subject to the restriction (7.7).

Although one could proceed by analytic geometry, or by the method of Lagrange multipliers, they both lead to the solution of a quartic equation. A simpler method seems to be the following. Make a transformation of (7.7) to the variables (u,v) via

$$(7.7.2) \qquad \begin{aligned} x &= \hat{b}_1 + u \cos \theta_0 - v \sin \theta_0 \\ y &= \hat{b}_2 + u \sin \theta_0 + v \cos \theta_0 \end{aligned}$$

where

$$\tan 2\theta_0 = \frac{2b}{c - a} .$$

Thus, we want to extremize the function of (u,v) obtained by substituting (7.7.2) into (7.7.1), subject to the restriction

$$a_1 u^2 + a_2 v^2 = 1,$$

where as in Equations (5.4.1), (5.4.2) of Reference [2] we have

$$a_1 = a \cos^2\theta_0 + b \sin 2\theta_0 + c \sin^2\theta_0$$

$$a_2 = a \sin^2\theta_0 - b \sin 2\theta_0 + c \cos^2\theta_0.$$

Now if we let $\cos \phi = \sqrt{a_1}\, u$, $\sin \phi = \sqrt{a_2}\, v$ and set

$$a_i = \frac{\cos \theta_0}{\sqrt{a_i}}\,, \qquad b_i = \frac{\sin \theta_0}{\sqrt{a_i}}\,, \qquad i = 1,2$$

our problem becomes that of extremizing the functions

$$(7.7.3) \qquad f(\phi) = \frac{b_1 \cos \phi + a_2 \sin \phi + \hat{b}_2}{a_1 \cos \phi - b_2 \sin \phi + \hat{b}_1}$$

and

$$(7.7.4) \quad g(\phi) = (a_1 \cos \phi - b_2 \sin \phi + \hat{b}_1)^2 + (b_1 \cos \phi + a_2 \sin \phi + \hat{b}_2)^2$$

for the range of values $0 < \phi \leq 2\pi$. This can be done by elementary techniques of calculus. One can show in a straightforward manner that

$$f'(\phi) = 0 \quad \text{if and only if}$$

$$(7.8) \qquad a + B_1 \sin \phi + B_2 \cos \phi = 0$$

where

$$a = b_2 b_1 + a_1 a_2$$

$$B_1 = \hat{b}_2 a_1 - \hat{b}_1 b_1$$

$$B_2 = \hat{b}_1 a_2 + \hat{b}_2 b_2$$

with capital letters those random variables which are functions of the estimates.

Now let Φ_0 be the unique random angle for which

$$\sin \Phi_0 = B_1/B, \qquad \cos \Phi_0 = B_2/B$$

where

$$B = \sqrt{B_1^2 + B_2^2} \, .$$

Then (7.8) becomes simply, by using the trigonometric identity for cosine of a sum,

$$\cos(\Phi_0 - \phi) = -a_1/B.$$

Letting Θ be the angle on $(0,\pi)$ such that $\Theta = \arccos(-a_1/B)$ we have, since cosine is an even function,

$$\Omega_1 = \Phi_0 - \Theta, \qquad \Omega_2 = \Phi_0 + \Theta$$

as the two solutions of (7.8). Thus, we have as the limits on the true bearing angle as in (7.5.3), the angles

$$\phi_1 = \arctan f(\Omega_1) \quad \phi_2 = \arctan f(\Omega_2).$$

This accomplishes the location of local extremum of f. Graphical examination of the function on $(0,2\pi)$ may be necessary to see if these are the true extremum.

Proceeding with the extremizing of g we see in a manner exactly similar for f that

$$g'(\phi) = 0 \qquad \text{if and only if}$$

(7.9) $a_1 \sin 2\phi + a_2 \cos 2\phi + C_1 \sin \phi + C_2 \cos \phi = 0$

where

$$2a_1 = \mathfrak{a}_1^2 - \mathfrak{a}_2^2 + \mathfrak{b}_1^b - \mathfrak{b}_2^2$$

$$a_2 = \mathfrak{a}_1 \mathfrak{b}_2 - \mathfrak{b}_1 \mathfrak{a}_2$$

$$C_1 = \hat{b}_1 \mathfrak{a}_1 + \hat{b}_2 \mathfrak{b}_1$$

$$C_2 = \hat{b}_1 \mathfrak{b}_2 - \hat{b}_2 \mathfrak{a}_2$$

and again upper case letters denote random variables. Unfortunately the solutions of (7.9), call them Γ_1, Γ_2, must be found in general by some numerical technique. However, this is easily accomplished by a computer so that we proceed no further.

Having found the solutions Γ_i i = 1,2 the limits on the velocity which were sought can be written

$$V_1 = \min(\sqrt{g(\Gamma_1)}, \qquad \sqrt{g(\Gamma_2)})$$

(7 9.1)

$$V_2 = \max(\sqrt{g(\Gamma_1)}, \qquad \sqrt{g(\Gamma_2)}).$$

In the special case that $\mathfrak{b} = 0$, (7.9) simplifies somewhat. This circumstance occurs whenever the observations are drawn symmetrically on each overflight, so that it is of practical interest.

In this case $\theta_0 = 0$, $a_1 = a$, $a_2 = c$

$$a_1 = \frac{1}{\sqrt{a}}, \qquad b_2 = 0 = b_1, \qquad a_2 = \frac{1}{\sqrt{c}}$$

$$2a_1 = \frac{1}{a} - \frac{1}{c}, \qquad a_2 = 0$$

$$\beta_1 = \hat{b}_1/\sqrt{a}, \qquad \beta_2 = -\hat{b}_2/\sqrt{c}.$$

Thus, the equation (7.9) reduces to

$$(7.10) \qquad \tfrac{1}{2}(\tfrac{1}{a} - \tfrac{1}{c}) \sin 2 + \frac{b_1}{\sqrt{a}} \sin = \frac{b_2}{\sqrt{c}} \cos \phi .$$

If we further specialize and assume that $a = c$, we see from (7.10) that the equation we now must solve is merely

$$(7.11) \qquad \tan \phi = \hat{b}_2/\hat{b}_1$$

which has solutions

$$\cos \phi = \pm \frac{\hat{b}_1}{V}, \qquad \sin \phi = \pm \frac{\hat{b}_2}{V}$$

where V is the estimate of velocity defined in (7.2). From (7.7.4) we see that $a = c$ implies that the function we must extremize is merely

$$g(\phi) = \left(\frac{\cos \phi}{\sqrt{a}} + \hat{b}_1 \right)^2 + \left(\frac{\sin \phi}{\sqrt{a}} + \hat{b}_2 \right)^2$$

and we obtain formally from (7.9.1) the answer (which was obvious to begin with)

$$V_2 = V + \frac{1}{\sqrt{a}} \quad , \quad V_1 = V - \frac{1}{\sqrt{a}}$$

VIII. DETERMINATION OF *A POSTERIORI* DISTRIBUTION OF HEADING AND VELOCITY

If the O. S. is making observations in an area which is dense with targets, such as a shipping lane or a point of concentration like the Straits of Gibralter, and the frequencies of all types of ships and their velocities over a recent period of time are known then we can make tabulations of the percentages of ships of various types proceeding at different velocities in certain headings. This characterization of the area we assume can be accomplished with an *a priori* distribution of speed and heading. The *a priori* density of speed $f_2(s)$ we take to be $N(s_0, d^2)$ with the mean speed s_0 known as is the standard deviation d. The *a priori* density of heading f_w is

$$f_w(h) = \begin{cases} f_0(h - w + 2) & -\pi < h < -\pi + w \\ f_0(h - w) & -\pi + w < h < \pi, \end{cases}$$

where

$$(8.1) \quad f_0(h) = p_1 \exp\left(\frac{-h^2}{2\sigma_1^2} \right) + p_2 \exp\left(\frac{h - \pi \operatorname{sgn} h}{2\sigma_2^2} \right) -\pi < h < \pi,$$

with p_i, σ_i, for $i = 1,2$, and the angle w known. For example, w is the angle the shipping lane makes with the direction of travel with the O. S.

In order to visualize the behavior of the density, we present a typical graph of (8.1):
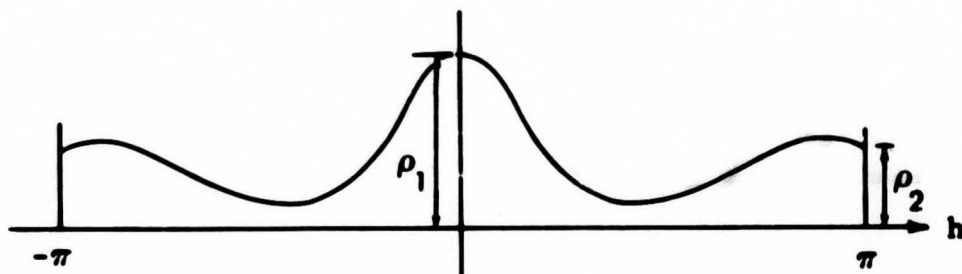

FIGURE 3

In order for (8.1) to be a *bona fide* density, we must have

$$\int_0^\pi f_0(h)dh = \tfrac{1}{2}.$$

By direct integration, this is equivalent to $p_1 > 0$ and

$$\tfrac{1}{2} = p_1\sigma_1\sqrt{2\pi}\left[\Phi\left(\frac{\pi}{\sigma_1}\right) - \tfrac{1}{2}\right] + p_2\sigma_2\sqrt{2\pi}\left[\tfrac{1}{2} - \Phi\left(\frac{-\pi}{\sigma_2}\right)\right].$$

Following the usual notation, Φ denotes the cumulative distribution function of the standard $N(0,1)$ variable which again violates our own convention that Greek letters are vectors.

Note that the angle w can be chosen so as to vary the heading of the greater percentage of ships away from or toward the direction of travel of the O. S. Further, the constants p_i can be chosen so

as to make the greater percentage of ships heading easterly, say, rather than westerly and of course either may be zero.

Thus, having obtained an observation of the random variable (Φ,V), call it (ϕ,v), we than may obtain via Bayes' theorem the *a posteriori* distribution of the heading and speed (h,s), call it q, namely

$$(8.2) \qquad q(h,s|\phi,v) = \frac{g(\phi,v|h,s)f_w(h)f_2(s)}{\displaystyle\int_0^\infty\int_0^{2\pi} f(\phi,v|h,s)f_w(h)f_2(s)\,dh\,ds} \, .$$

This integration can, in general, only be carried out by machine computation and, moreover, is subject to all the objections that one can raise concerning the use of prior distributions. However, in certain special cases the density (8.2) could then be used to study the distribution of the true heading and the true speed, given the observations, with meaningful interpretation of the results.

ACKNOWLEDGEMENTS

I would like to express my appreciation to Jack Israel for introducing me to this problem and for pointing out those facets which were of practical importance.

I would like to thank David L. Johnson for reading an early draft of this manuscript and offering suggestions about explanation which has materially helped the exposition. Also, I want to thank Joel Frenk who in programming for computation detected an error in the first version of this document.

BIBLIOGRAPHY

[1] Anderson, T. W., (1958), "Introduction to Multivariate Statistical Analysis," John Wiley and Sons.

[2] Saunders, S. C. and Johnson, D. L., (1964), "On the Estimation of Location Coordinates from Data Taken from a Manned Orbital Laboratory," Boeing Document D1-82-0320.